Assignment 7

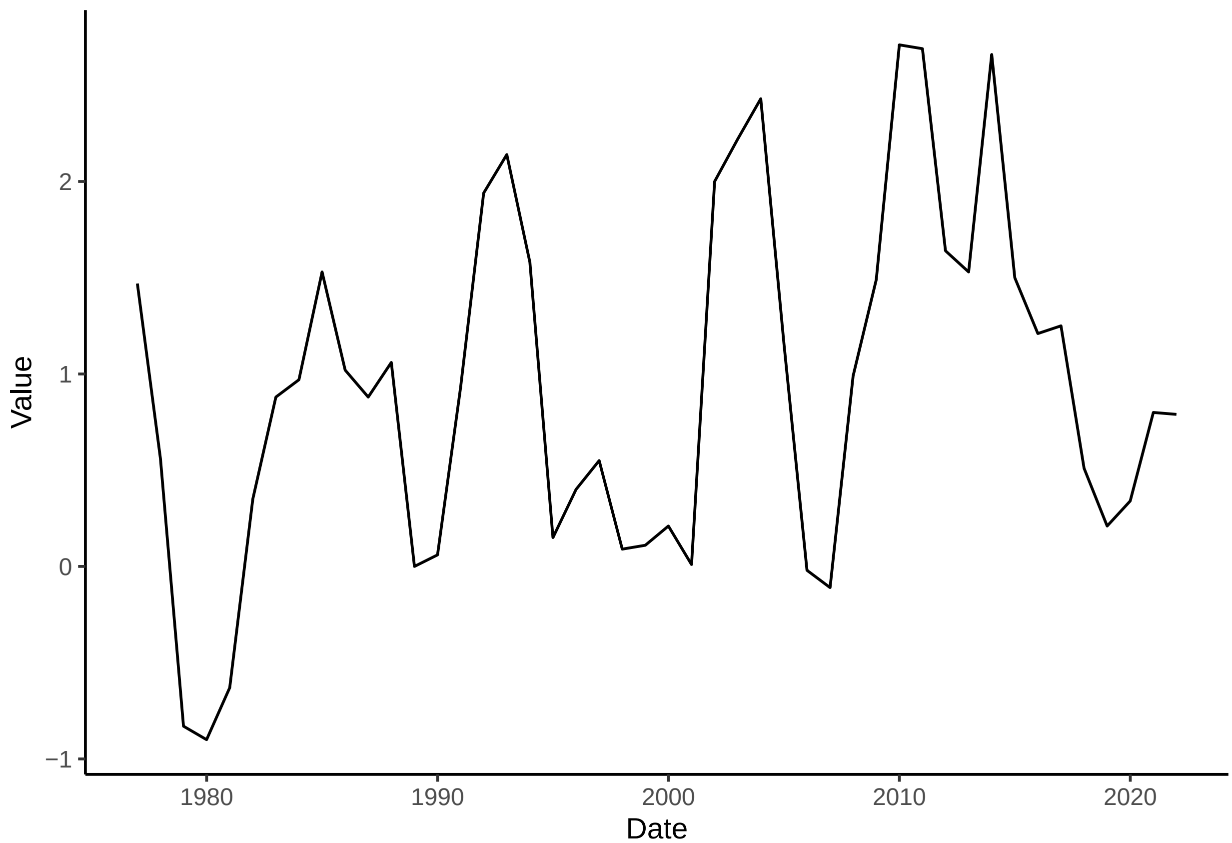
By Ben Geiser and Gage Brockington

2022-03-23

First, I will insert the necessary libraries and the datasets.

```
#load libraries
library(tidyverse)
library(ggplot2)
library(quantmod)
library(Quandl)

#read in the data from fred
#first data set is the real GPD growth
#second data set loaded in will be:
#10-Year Treasury Constant Maturity Minus 2-Year Treasury Constant Maturity
Quandl.api_key("Bsjtos3Rumnys9NRFspt")
macro_data <- Quandl("FRED/GDPC1",
                     start_date="1976-06-01",
                     end_date="2021-12-31",
                     collapse = "annual",
                     transform = "rdiff")
macro_data1 <- Quandl("FRED/T10Y2Y",
                      start_date="1976-06-01",
                      end_date="2021-12-31",
                      collapse = "annual")
macro_data<-data.frame(macro_data)
macro_data1<-data.frame(macro_data1)
#delete NA rows
#macro_data1<- macro_data1[-c(182),]

GDP1<- macro_data %>% select(
  Date, Value
)
yield<- macro_data1 %>% select(
  Date, Value
)

#plot GDP and the yield
GDP_graph<-ggplot()+
  geom_line(data = GDP1, aes(x = Date, y = Value))+
  theme_classic()

print(GDP_graph)
```


```
yield_graph<-ggplot()+
  geom_line(data = yield, aes(x = Date, y = Value))+
  theme_classic()

print(yield_graph)
```



After I have plotted the GDP and the yield curve, I can create a simple regression model and summarize the output.

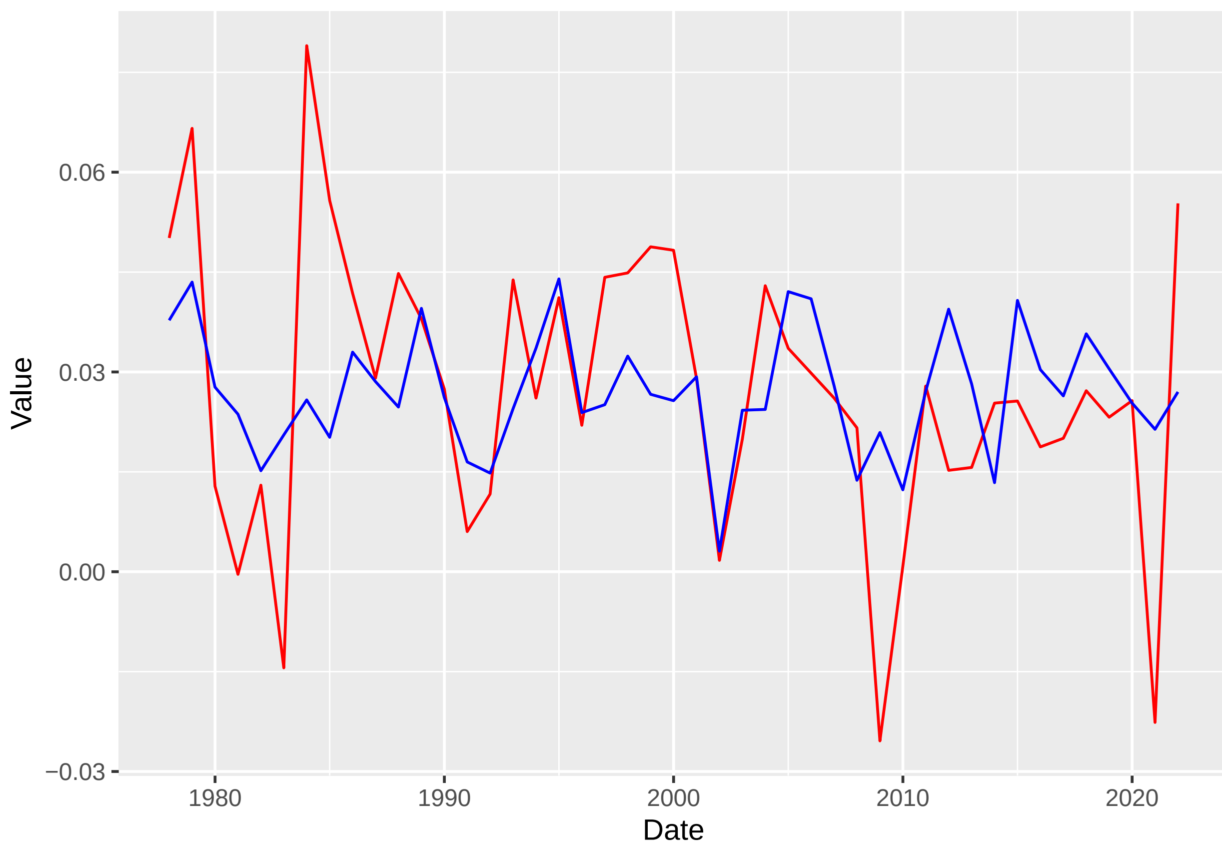
```
#combine the two dataframes before analysis

diff_int_spread<- diff(as.numeric(unlist(yield$Value)))
macro_data$lagged<- diff_int_spread
fullmodel <- lm(Value~lagged, data = macro_data)
summary(fullmodel)
```

```
##
## Call:
## lm(formula = Value ~ lagged, data = macro_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.04631 -0.01046 -0.00187  0.01236  0.05319
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.026883   0.002915   9.222  9.5e-12 ***
## lagged      0.011953   0.003906   3.060   0.0038 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01955 on 43 degrees of freedom
## Multiple R-squared:  0.1789, Adjusted R-squared:  0.1598
```

```
## F-statistic: 9.366 on 1 and 43 DF,  p-value: 0.003799
```

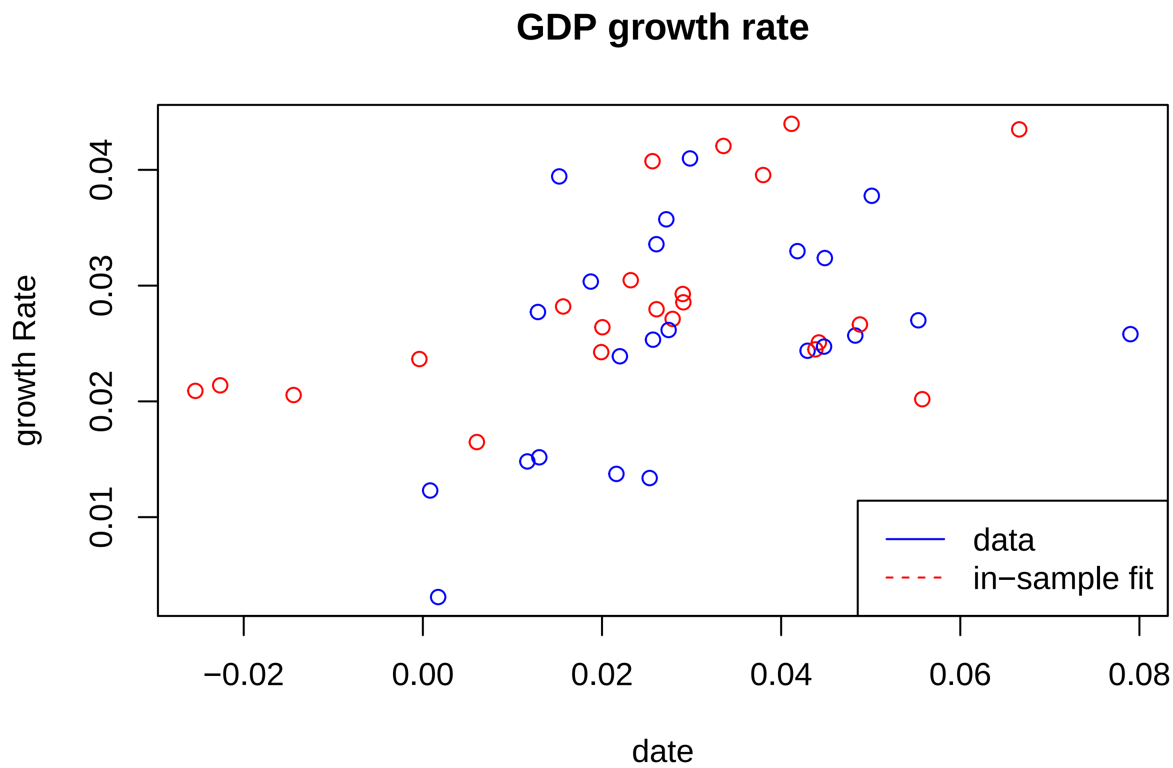
Based on the above regression output, we can analyze the results in both a statistical and economic perspective. Statistically, for every 1% increase in the yield curve, there will be a 0.011953% increase in the real GDP. Economically, this makes sense as the increase in the value of a bond should correlate to a higher GDP. When there is an inverted yield curve, this should signify that a recession is on the horizon. Now, I will examine the in-sample fit.

```
fit=ts(fullmodel$fitted.values)
fit<- data.frame(fit)
GDP1$fit <- fit
GDP1$fit<- unlist(GDP1$fit)

data_combine=cbind(GDP1$Value,fit)

ggplot()+
  geom_line(data = GDP1, aes(x=Date, y = Value), color="red")+
  geom_line(data = GDP1, aes(x=Date, y = fit), color="blue")
```




```
plot(data_combine,main="GDP growth rate",
     plot.type=c("single"),
     xlab = "date",
     ylab="growth Rate",
     col=c("blue","red"),
     lty=1:2)
```

```
legend("bottomright",
       legend=c("data","in-sample fit"),
       col=c("blue","red"),
       lty=1:2)
```



3.) The in-sample fit shows that the model is an okay fit for predicting GDP. The adjusted R^2 is 0.1598, which means that 15% of the variance in GDP is predicted by the yield curve.

4.) If we were to use a model that was made ten years ago, it would not perform very well due to the impossibility of predicting the COVID recession. A sample mean would not perform as well as the model, due to the predictions being thrown off by the variances.

5.) Under the conditions in the previous question, the model would perform better than the sample mean, however, the sample mean is still useful as it is isolated from the booms and busts of the economy.